

РЕЦЕНЗИЈА

НА ДОКТОРСКАТА ДИСЕРТАЦИЈА ПОД НАСЛОВ „МОДЕЛ НА ПРЕДВИДУВАЊЕ НА ОПШТИТЕ ИЗБОРИ ВО КОСОВО СО КОРИСТЕЊЕ НА ТЕХНИКАТА ЗА МАШИНСКО УЧЕЊЕ KNEIGHBORSCCLASSIFIER“, ИЗРАБОТЕНА ОД М-Р ЛАВДИМ БЕКИРИ, ПРИЈАВЕНА НА ФАКУЛТЕТ ЗА ИНФОРМАТИКА ПРИ УНИВЕРЗИТЕТ „ГОЦЕ ДЕЛЧЕВ“ ВО ШТИП

Со Одлука број 0206-841/2 од 26.12.2022 година донесена на 72. седница (електронска) на Наставно-научниот совет на докторски студии на Кампус 2 (биотехнички науки, техничко-технолошки и природно-математички науки) при Универзитет „Гоце Делчев“ во Штип, одржана на 26.12.2022 година, формирана е Комисија за оценка и одбрана на докторската дисертација под наслов „Модел на предвидување на општите избори во Косово со користење на техниката за машинско учење Kneighborsclassifier“, пријавена и изработена од кандидатот м-р Лавдим Бекири, во состав:

- проф. д-р Александар Крстев, претседател,
- проф. д-р Андон Мајхошев, член,
- проф. д-р Иван Чорбев, член,
- проф. д-р Мајлинда Фетаи, екстерен ментор, член,
- проф. д-р Зоран Здравев, интерен ментор, член.

Комисијата во наведениот состав, по прегледувањето на докторската дисертација, го поднесува следниов

ИЗВЕШТАЈ

Анализа на трудот

Докторската дисертација со наслов „Модел на предвидување на општите избори во Косово со користење на техниката за машинско учење kneighborsclassifier“, пријавена и изработена од м-р Лавдим Бекири, е напишана на македонски јазик, со правилно користење на научна и стручна терминологија, на 155 страници, на А4 формат, согласно со интерните акти на Универзитетот за пишување на дипломски, магистерски и докторски трудови. Резултатите се прикажани со 22 слики, 52 табели.

Дисертацијата ги опфаќа следниве поглавја: Вовед, Преглед на литературата, Цели на истражувањето, Методи на истражувачката работа, Идентификување на фактори за мерење, Резултати, Осмислување на модел за предвидување на општите избори во Косово со помош на техниката за машинско учење KNeighborsClassifier, Наоди и резултати, Статистичка анализа на податоци за предвидување на изборите во Косово, Заклучок, Користена литература.

Вреднување на поединечни делови од докторската дисертација

Вовед - во првото поглавје е дадена кратка анализа и осврт на постоечките трудови од областа и кратки објаснувања на основните поими како: предикција на избори, анализа на фактори на информации за гласање и разни социјални медиумски платформи со нивна примена. Во овој дел се презентирани целите на истражувањето, главните придобивки, како и краткиот преглед на содржината.

Преглед на литературата - вториот дел дава преглед на литературата и анализа на основните пристапи и видови на предикции на предвидување на избори и предвидување на резултатот од изборите со користење на софтверски системи, како и таксономијата со кои истите се класифицираат.

Цели на истражувањата - третиот дел е посветен на целите, со објаснување на целите и се користат наодите ги идентификуваат различните методи кои се користат, предностите и недостатоците на различните пристапи и методите што се користат во моментот и кои покажале најефикасни резултати и дадени се препораки.

Во следните три дела се презентирани главните резултати на докторската дисертација и тие воедно го претставуваат нејзиниот најважен дел.

Методи на истражувачката работа - четвртиот дел ја опишува користената методологија на истражување, која е квантитативна методологија на истражување. Квантитативна методологија на истражување во адресирање на емпириските аспекти на податоците и извлекување заклучоци и стекнување знаење од анализирани податоци и користени алгоритми за вештачка интелигенција за предвидување на изборите.

Методологијата на истражување опфаќа фундаментално истражување кое опфаќаше анализи на различни алгоритми за машинско учење и анкети за изборите. Потоа користевме применети истражувања за да ја тестираме нашата хипотеза. Главната цел на истражувањето беше да се истражат перформансите на алгоритмите за машинско учење за предвидување избори и можностите на овој пристап и неговите примени во подобрувањето на процесот на прогнозирање на изборите. Ја користевме Адаптивната хеуристичка алатка и го смисливме алгоритамскиот модел за најважните проблеми (MIP) за да го предвидиме исходот од парламентарните избори на Косово во 2021 година. Адаптивниот хеуристички алатник е репертоар на правила или хеуристики достапни за еден вид во даден момент од неговата еволуција [1]. Еден од најосновните алгоритми во адаптивниот хеуристички алатник и оној кој е во фокусот на овој напис е стратегијата ТТВ (take-the-best). [1] Тврдеше дека донесуваме одлуки преку избор и примена на различни брзи и штедливи хеуристики содржани во нашата когнитивна „приспособлива кутија со алатки“.

Идентификување на фактори за мерење - петтиот дел го опишува идентификувањето на фактори за мерење. Теоретски, алгоритмите за учење можат да помогнат да се подобрат резултатите од анализата и да се генерираат силни статистички модели кои можат да се користат за предвидување на други видови податоци, како што се цените на акциите. Во пракса, тешко е да се добие соодветен модел за да се изврши корисна предвидувачка анализа за одредена ситуација. Со цел да се намали овој проблем, обично се користат статистички модели за да се опише статистичкиот процес наречен латентна логика. Во овој модел, движењето на агентот е претставено со латентни променливи. Овие променливи се поврзани со движењето, однесувањето и одговорот на агентот, но не се дел од агентот. Овој метод е многу популарен при моделирање на геополитички ситуации со сложена динамика во системот. Ова е методот што го користиме за да ја разбереме динамиката на однесувањето на гласањето.

Резултати - шестиот дел ја предвидува излезноста на општите избори во 2021 година во Косово. Користен е надгледуван алгоритам за машинско учење заснован на одредени карактеристики на поединци и нивната историја на гласање. Го користи програмскиот јазик Python. За подобро разбирање на податоците се претпоставува дека волјата на луѓето да бидат вклучени во општите и примарните избори се разликувала. Затоа, за да се предвиди нивото на излезност на општите избори во 2021 година го обучивме моделот користејќи ги минатите податоци за излезноста на општите избори. Историјата на гласање, на пример, „pr00g“, доаѓа како стринг променлива наместо цел број вредност. При иницијализирање на атрибутите на класата „Гласач“, полињата со префикс „pr“ беа претворени во цели броеви. Дополнително, вредностите на преостанатите полиња како што се брачниот статус, етничката припадност и професионалната индустрија се претставени како цели броеви. Сепак, неколку ќелии под полињата „возраст“ и „предел“ беа празни, така што беше доделено „0“ за тие празни ќелии за да може да се изврши програмата.

Осмислување на модел за предвидување на општите избори во КОСОВО со помош на техниката за машинско учење KNeighborsClassifier, седмиот дел го опишува моделот кој е развиен. Со оглед на фактот дека програмата треба да предвиди дали поединецот ќе се појави на избирачко место или не, методот на машинско учење најдобро одговара за овој модел. Се користи библиотека Scikit-learn, бидејќи библиотеката се состои од различни алгоритми за класификација и регресија. Откако пробав различни техники за машинско учење, забележав дека методите KNeighbors и Decision Tree даваат попрецизни предвидувања од другите класификатори. Го избрав KNeighbors, надгледувана метода за машинско учење, бидејќи KNeighbors ги анализира соседите „n“ околу избраниот заговор на податоци. Поконкретно, во мојот модел, алгоритмот KNeighbors (n=10) ги испитува

десетте најблиски соседи на избраниот гласач. Секоја од овие податочни парцели претставува еден гласач. Користејќи го класификаторот KNeighbors, алгоритмот ги учи шемите на историјата на гласање на овие гласачи за да ја предвиди излезноста на гласачите за општите избори во 2021 година.

Наоди и резултати, во осмиот дел се дадени наодите и резултатите од истражувањето кои се публикувани како научни трудови во меѓународна IEEE конференција.

Статистичка анализа на податоци за предвидување на изборите во Косово - во деветтиот дел кандидатот се осврнува на статистичките анализи од истражувањето.

Заклучок – ова поглавје ги сублимира сите постигнати резултати со оваа докторска дисертација и изведува јасни, прецизни и правилни заклучоци. Дополнително, поглавјето дава насоки за идни истражувања.

Користена литература – кандидатот цитирал 97 референци, од кои најголемиот дел се научни трудови објавени во реномирани и релевантни научни списанија или во зборници од реномирани научни конференции.

Предмет на докторската дисертација

Предмет на докторската дисертација „Модел на предвидување на општите избори во Косово со користење на техниката за машинско учење kneighborsclassifier“, на кандидатот м-р Лавдим Селајдин Беќири дава сеопфатна анализа на можностите за предвидување на избори и предвидување на резултатот од изборите со користење на софтверски системи. Факторите кои ги мотивираат истражувачките институции и поединците да го земат предвид влијанието на истражувањето врз предвидувањето на изборите се многубројни. Разбирањето на влијанието на различни софтверски алатки, алгоритми и софтверски апликации за социјално вмрежување врз предвидувањето на изборите е витален, и честопати занемарен, елемент за прогнозирање на изборните резултати. Прегледот на литературата беше спроведен за да се испитаат методите и тековните софтверски апликации и практики, како и проекти за предвидувања на изборите. Во рамките на истражувањето се користеше надгледуван алгоритам за машинско учење заснован на одредени карактеристики на поединци и нивната историја на гласање. За експерименталната евалуација направена е имплементација на класификаторот KNeighbors кој даде попрецизни предвидувања од другите класификатори.

Дополнителен придонес на докторската дисертација е нов алгоритам алгоритамски модел за најважниот проблем (МИП) за да го предвидиме исходот од парламентарните избори.

Темата е од исклучително научна и апликативна важност, модерна и доста актуелна, посебно ако се земе предвид постојаното зголемување на потребата за демократизација на институциите за приклучок кон Европската Унија.

Цели и нивна разработка

Целите на истражување во оваа докторска дисертација се:

- Откривање на нов модел за предвидување на победникот на парламентарните избори;
- Имплементација и експериментална евалуација на некои од предложените фактори за предвидување;
- Испитување на најважните фактори кои влијаат на гласањето;
- Испитување на тие фактори со користење на разни алгоритми за машинско учење;
- Испитување на моделот кој е развиен;
- Статистичка анализа на податоците кои се добиени и пресметување на маргините на грешка;
- Предлог на нов модел за предикција на резултатите од изборите.

Овие цели во целост се реализирани со извршени експерименти, соодветно добиени и интерпретирани резултати и изведени правилни заклучоци.

Научен придонес

Оригиналниот научниот придонес на докторската дисертација се состои во извршената анализа на идентификување на различните методи кои се користат, предностите и недостатоците на различните пристапи и методите што се користат во моментот, кои покажале најефикасни резултати и се дадени препораки. Развиен е нов метод за препознавање на главните фактори кои влијаат на изборите. Прво е користена адаптивна хеуристика. Со цел да се осмислат и да се проценат факторите кои влијаат, смислен е алгоритамски модел за најважниот проблем (МИП) за да се предвиди исходот од парламентарните избори на Косово во 2021 година и е заснована на стратегијата ТТБ (земи го најдоброто). Користена е анализа на пристапот на предвидување на изборите и метриката на перформанси (варијанса) со користење на алгоритам за машинско учење за длабоко учење. Податоците се изведени од оригинално собрана база на податоци од анкета која ги содржи факторите на влијание кои претходно беа идентификувани и оценети во однос на парламентарните избори во Косово во февруари 2021 година, се реализирани. Во рамките на истражувањето се користеше надгледуван алгоритам за машинско учење заснован на одредени карактеристики на поединци и нивната историја на гласање. Во текот на студијата се истражуваат специфичните фактори кои влијаат на изборниот процес во Косово. Користена е вкрстена валидација за да пресметаме вкупно 200 прогнози од примерокот, по една за секој од градовите и партиите, во последните 100 дена од општите избори од февруари 2021 година. По неколку експерименти дојдено е до заклучок дека класификаторот KNeighbors даде попрецизни предвидувања од другите класификатори. Затоа е избран KNeighbors класификатор како надгледуван метод за машинско учење, бидејќи KNeighbors ги анализира соседите „n“ околу избраниот заговор на податоци.

За експерименталната евалуација е направена имплементација на класификаторот KNeighbors кој даде попрецизни предвидувања од другите класификатори. Дополнителен придонес на докторската дисертација е нов алгоритам алгоритамски модел за најважниот проблем (МИП) за да се предвиди исходот од парламентарните избори.

Докторската дисертација изработена под менторство на проф. д-р Зоран Здравев, редовен професор на Факултетот за информатика при Универзитет „Гоце Делчев“ во Штип, како интересен ментор и проф. д-р Мајлинда Фетаји, редовен професор на Факултетот за современи науки и технологии, Департмент по компјутерски науки при Универзитет на Југоисточна Европа во Тетово, како екстерен ментор, е успешно завршена. Трудот е оригинален и јасно напишан, а целокупната содржина заедно со добиените резултати и изведените заклучоци е јасно и концизно презентирани. Докторскиот труд има како научно, така и апликативно значење, бидејќи предложените противмерки може да се имплементираат во иднина.

ЗАКЛУЧОК И ПРЕДЛОГ

Комисијата за оценка и одбрана детално ја разгледа докторската дисертација со наслов „Модел на предвидување на општите избори во Косово со користење на техниката за машинско учење Kneighborsclassifier” на кандидатот м-р Лавдим Бекири и донесе заклучок дека истата претставува оригинален, самостоен, прецизно дефиниран, јасно оформен научен труд со систематски разработена проблематика и оригинални научни истражувања и резултати. Докторската дисертација врз основа на содржината, обемот и постигнатото ниво на квалитет на научна работа ги задоволува и исполнува условите потребни за изработка на докторска дисертација.

Потврда за научната тежина на темата се и објавените оригинални научни трудови од дисертацијата во реномирани меѓународни научни списанија и конференции.

Врз основа на претходно изнесеното, Комисијата има чест да му предложи на **Наставно-научниот совет на докторски студии на Кампус 2 да ја прифати позитивната рецензија на докторската дисертација со наслов „Модел на предвидување на општите избори во Косово со користење на техниката за машинско учење Kneighborsclassifier” на кандидатот м-р Лавдим Бекири и да одобри јавна одбрана на истата.**

РЕЦЕНЗЕНТСКА КОМИСИЈА

Д-р Александар Крстев,
вонреден професор на Факултет за информатика,
Универзитет „Гоце Делчев“ – Штип, претседател, с.р.

Д-р Зоран Здравев,
редовен професор на Факултет за информатика, Универзитет
„Гоце Делчев“ – Штип, интерен ментор, член, с.р.

Д-р Мајлинда Фетаји,
редовен професор на Факултет за современи науки и
технологии, Универзитет на Југоисточна Европа – Тетово,
екстерен ментор, член, с.р.

Д-р Иван Чорбев,
редовен професор на Факултет за информатички науки и
компјутерско инженерство, УКИМ – Скопје, член, с.р.

Проф. д-р Андон Мајхошев,
редовен професор на Правен факултет, Универзитет „Гоце
Делчев“ – Штип, интерен ментор, член, с.р.